



Tvrtko Tadić, Microsoft, Redmond, USA

SEZONALNOST PODATAKA I POMIČNI PROSJEK

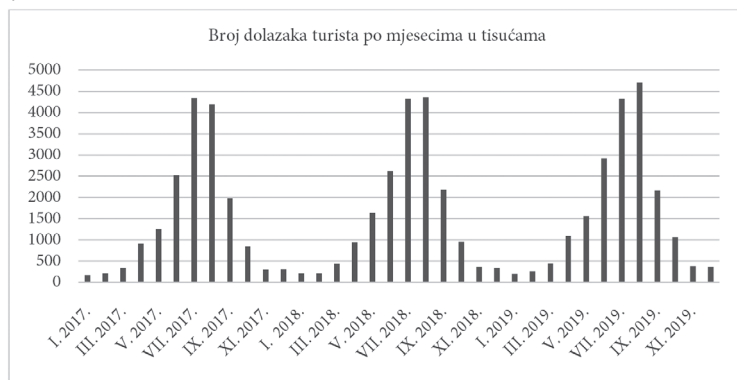
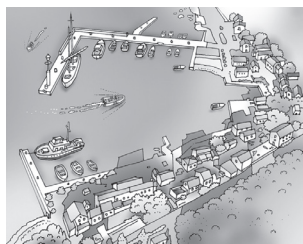
Ljudi često vole bilježiti vrijednosti nekih svojstava ili pojava kroz vrijeme, kao što su masa osobe, iznos novca na računu, temperatura... Tako prikupljene podatke zovemo **vremenskim nizom**.

Kod takvih mjerenja obično se odredi koliko će se često ono provoditi. To će ovisiti o potrebi. Tako masu osobe možemo mjeriti svaki dan, iznos novca na računu svaki mjesec, temperaturu svaki sat, a popis stanovništva se, zbog zahtjevnosti njegova provođenja, radi svakih 10 godina.

Cilj ovako prikupljenih podataka je usporediti **trenutnu vrijednost s vrijednostima zabilježenima prije toga trenutka**. U većini slučajeva radi se usporedba trenutne vrijednosti s vrijednošću zabilježenom u prethodnome mjerenju. Tako bi svi ovaj mjesec rado vidjeli više novca na računu nego što je to bilo mjesec prije. No, kao što ćemo vidjeti, to neće uvijek imati smisla.

Sezonalnost podataka

Kao što Matkači znaju, Hrvatska je popularno ljetno turističko odredište. Državni zavod za statistiku mjesečno bilježi broj turističkih dolazaka tijekom cijele godine¹. Na Slici 1. možemo vidjeti podatke za razdoblje od 2017. do 2019. godine².



Slika 1. Mjesečni broj dolazaka turista

¹Sa stranica Državnog zavoda za statistiku www.dzs.hr preuzet je skup podataka Dolasci i noćenja turista, rujan 2023.

²Izabrali smo podatke prije 2020. godine jer je pandemija korona virusa poremetila dinamiku dolaska turista.



Ako usporedimo podatke koliko je (u postotcima) broj turista veći ili manji u odnosu na prethodni mjesec, tj. ako primijenimo formulu:

$$\left(\frac{\text{broj turističkih dolazaka u mjesecu}}{\text{broj turističkih dolazaka u prethodnom mjesecu}} - 1 \right) \times 100 \%,$$

dobit ćemo brojeve koji će biti nejasni za tumačenje. Pogledamo li tablicu 1., ne možemo odmah vidjeti povećava li se ili smanjuje broj turista koji dolaze u Hrvatsku.

	2018.	2019.
Siječanj	-32.8%	-40.4%
Veljača	3.8%	26.9%
Ožujak	100.9%	70.5%
Travanj	116.1%	145.8%
Svibanj	72.8%	41.9%
Lipanj	60.0%	86.2%
Srpanj	64.8%	48.1%
Kolovoz	0.9%	8.9%
Rujan	-49.9%	-53.9%
Listopad	-56.0%	-50.4%
Studeni	-61.4%	-63.7%
Prosinac	-6.2%	-4.9%

Tablica 1. Usporedba dolazaka turista u mjesecu s prethodnim mjesecom

	2018.	2019.
Siječanj	19.2%	-1.4%
Veljača	4.3%	20.5%
Ožujak	28.3%	2.3%
Travanj	3.8%	16.3%
Svibanj	30.7%	-4.5%
Lipanj	4.0%	11.1%
Srpanj	-0.4%	-0.1%
Kolovoz	3.9%	7.9%
Rujan	10.7%	-0.9%
Listopad	13.3%	11.7%
Studeni	20.8%	5.1%
Prosinac	11.1%	6.6%

Tablica 2. Usporedba dolazaka turista u mjesecu s istim mjesecom prethodne godine

Na temelju pregleda Slike 1. i poznavanja turizma u Hrvatskoj, možemo zaključiti:

- Najviše turista dolazi u ljetnim mjesecima.
- Najviše turista dolazi dok traju školski praznici, dakle u srpnju i kolovozu.

Općenito, možemo pretpostaviti da dolazak turista ovisi o tome u kojem se razdoblju u kalendarskoj godini nalazimo. Ljudi putuju ovisno o vremenskim prilikama, blagdanima, školskim praznicima... Zato umjesto usporedbe s prethodnim mjesecom više smisla ima uspoređivati pojedini mjesec s istim mjesecom u prethodnoj godini:

$$\left(\frac{\text{broj turističkih dolazaka u mjesecu}}{\text{broj turističkih dolazaka u prethodnom mjesecu}} - 1 \right) \times 100 \%$$

Iz tablice 2. može se naslutiti, barem za glavne mjesece, da broj dolazaka raste.



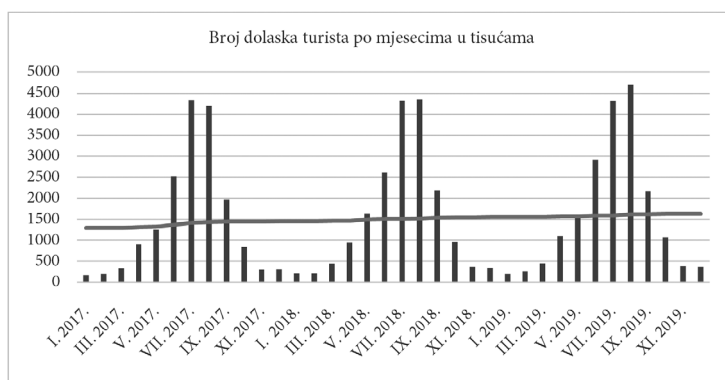


Pomični prosjek

Kako bismo imali jasniju sliku trenda, razvijena je ideja **pomičnog prosjeka** kao metode *desezonalizacije* podataka. Sa Slike 1. i iz razmatranja pretpostavimo da je *period sezonalnosti* vrijeme od jedne godine, tj. od 12 mjeseci. Zato definiramo pomični prosjek na sljedeći način:

broj turističkih dolazaka u intervalu $\left[\text{mjesec, isti mjesec prethodne godine} \right)$.

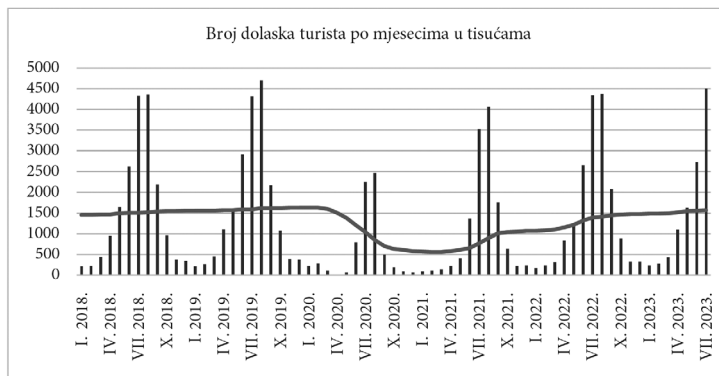
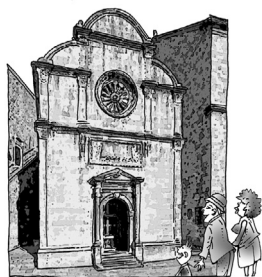
12



Slika 2. Originalni podatci s pomičnim prosjekom

Pogledajmo ponovo podatke ilustrirane na Slici 1. Za svaki mjesec izračunajmo vrijednost pomičnog prosjeka, označimo tu vrijednost točkom na grafu, te dužinama spojimo točke koje predstavljaju uzastopne vrijednosti. Rezultat je vidljiv na Slici 2.

Sa Slike 2. možemo vidjeti da je u razdoblju od 2017. do 2019. godine bio prisutan trend rasta broja dolazaka turista. Pogledajmo sada puno razdoblje od 2017. godine do srpnja 2023.



Slika 3. Podatci o dolasku turista s pomičnim prosjekom



Na Slici 3. možemo vidjeti kako je tijekom pandemije korona virusa 2020. godine došlo do pada broja dolazaka turista, da je oporavak počeo 2021. i 2022. godine, a da je 2023. ta brojka na putu da se vrati na prijašnje razine.

Zaključak

Ovdje smo na primjeru pokazali što je sezonalnost podataka i kako možemo analizirati takve podatke. Postoje i brojni drugi primjeri u kojima se može uočiti sezonalnost:

- Kad kompanije na burzi objavljuju tromjesečna (kvartalna) izvješća, uspoređuju trenutno tromjesečje s istim tromjesečjem prethodne godine.
- Potrošnja električne energije ima više elemenata sezonalnosti – struja se drukčije troši u različitim razdobljima dana, različitim dijelovima tjedna (radni dan u usporedbi s neradnim), te godišnjim dobima.
- ...

Sve ovo zahtijeva podatke i metode kako bi se moglo ustanoviti kakvi su trendovi. Pomični prosjek³ kao najosnovniji alat upoznali smo u ovome članku. Na temelju onoga što želimo izmjeriti, pozadine i vizualnog pregleda podataka, pretpostavljamo sezonalnost s periodom d i za trenutak t računamo pomični prosjek kao

$$\frac{\text{zbroj vrijednosti izmjerenih u intervalu } [t, t-d]}{d}$$

Matkači mogu sami pokušati naći neke zanimljive sezonalne podatke i ilustrirati kako ova metoda pomičnog prosjeka radi.

Podatci i pomični prosjek u Excelu

Excel⁴ tablice koje su korištene za izradu podataka u ovome članku (s poveznicom na izvorne podatke Državnog zavoda za statistiku) mogu se naći na stranici <https://web.math.hr/~tvrko/Matka/Turisti/>.

Kratko ćemo opisati kako smo napravili Sliku 3. koristeći proračunske tablice u Excelu.

Prvo smo sve mjesečne podatke o dolasku broja turista od 2017.⁵ do srpnja 2023. stavili na jednu stranicu tako da se (vidi Sliku 4.):

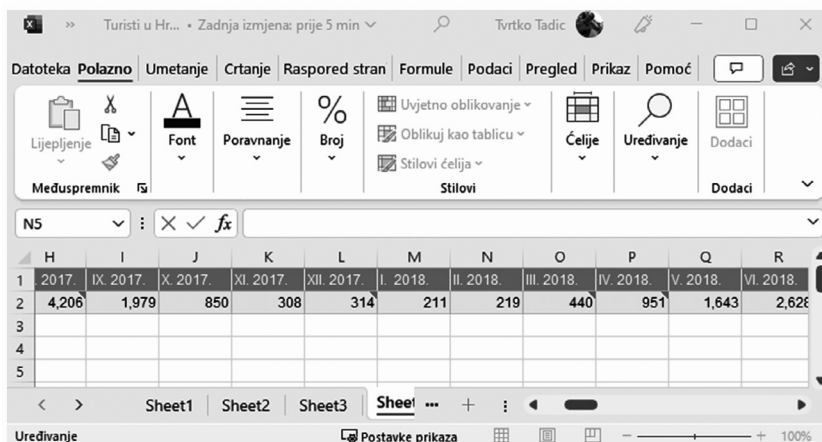
- u gornjem redu nalazi oznaka mjeseca i godine;
- u donjem redu broj dolazaka u tisućama.

³Pomični prosjek može se definirati na više načina. Ovdje je prikazan jedan.

⁴Excel je dio Office 365 paketa je koji je besplatno dostupan djelatnicima i polaznicima obrazovnog sustava u Hrvatskoj. Za više informacija posjetite <https://office365.skole.hr/>.

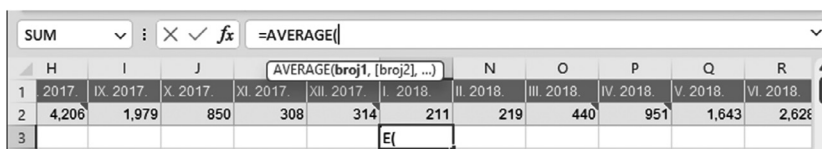
⁵Za izračun pomičnog prosjeka u prvim mjesecima 2018. godine trebaju nam podatci iz 2017. godine.



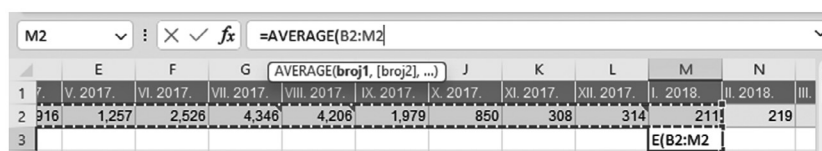


Slika 4. Unos broja dolazaka turista u Excelu

U sljedećem koraku ispod podatka za siječanj 2018. godine upišite =AVERAGE (i odaberite sva polja koja uključuju podatke od veljače 2017. do siječnja 2018. godine. Zatvorite zagradu i potvrdite unos (stiskom tipke Enter).

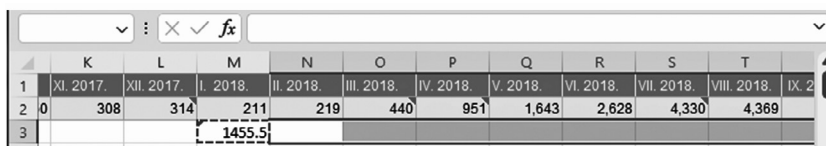


Slika 5. Izračun prosjeka



Slika 6. Odabir ćelija za izračun prosjeka

U ćeliji ispod podatka za siječanj 2018. godine sada će biti vrijednost pomičnog prosjeka za taj mjesec. Nakon toga ćemo izračunati tu vrijednost za sve ostale mjesece. Odaberimo tu ćeliju i napravimo copy (Ctrl+C).



Slika 7. Odabir ćelija u kojima ćemo izračunati pomični prosjek

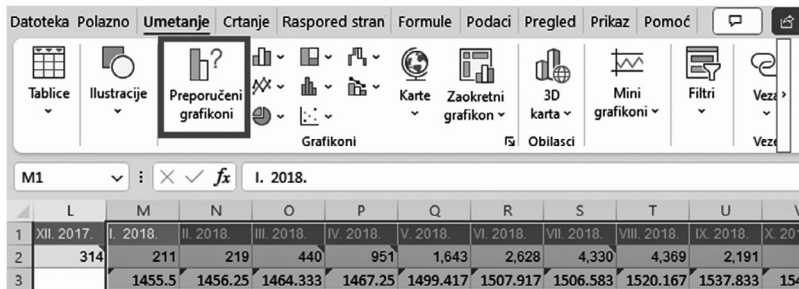


Odaberimo ćelije koje se nalaze desno od one za siječanj 2018. i napravimo paste (Ctrl+P). Sada smo dobili sve vrijednosti pomičnog prosjeka za željene mjesece.

	K	L	M	N	O	P	Q	R	S	T	U
1	.2017.	XII. 2017.	I. 2018.	II. 2018.	III. 2018.	IV. 2018.	V. 2018.	VI. 2018.	VII. 2018.	VIII. 2018.	IX. 2018.
2		308	314	211	219	440	951	1.643	2.628	4.330	4.369
3			1455.5	1456.25	1464.333	1467.25	1499.417	1507.917	1506.583	1520.167	1537.833

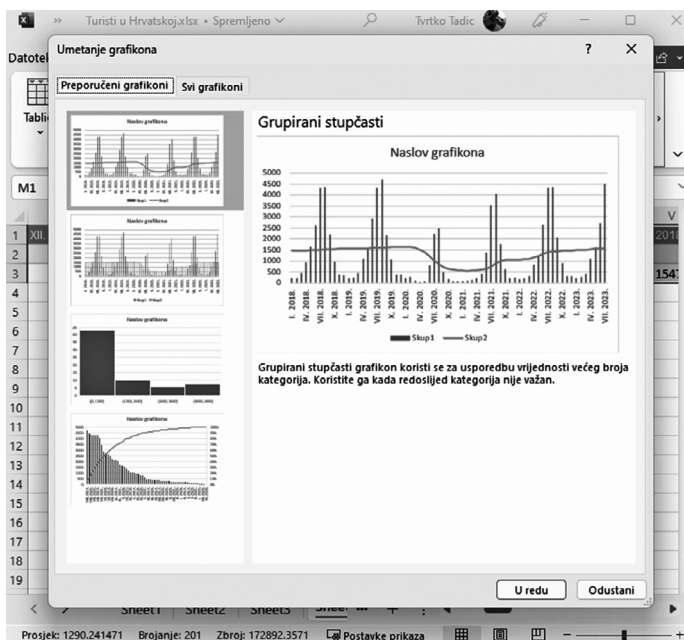
Slika 8. Izračun pomičnih prosjeka za sve željene mjesece

Preostaje nam nacrtati frekvencijski dijagram i *krivulju* pomičnog prosjeka. Odaberimo stupce koji pripadaju odgovarajućim mjesecima s nazivima i podacima. Na izborniku odaberimo Umetanje i tablicu Preporučeni grafikon.



Slika 9. Odabir podatka i izbor načina na koji stvaramo grafikon

Odaberemo izbor koji nam nudi grafikon koji imamo na Slici 10.



Slika 10. Izbor željenog tipa grafikona

